(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 387 876 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: **90104922.1**

(51) Int. Cl.5: **G06F 15/38**

(22) Date of filing: **15.03.90**

The title of the invention has been amended (Guidelines for Examination in the EPO, A-III, 7.3).

(30) Priority: **15.03.89 JP 63310/89**

(43) Date of publication of application: **19.09.90 Bulletin 90/38**

(84) Designated Contracting States: **DE FR GB**

(71) Applicant: **KABUSHIKI KAISHA TOSHIBA**
72, Horikawa-Cho Saiwai-ku
Kawasaki-shi Kanagawa-ken(JP)

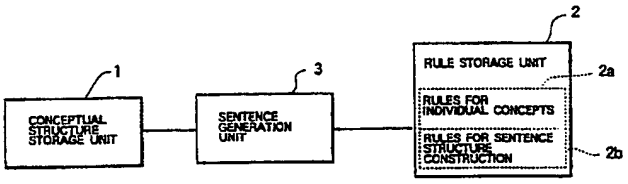
(72) Inventor: **Yoshimura, Yumiko,**
Toshiba-Shimomoge-Kazoku apart.
512 Shinmonoge, Takatsu-ku
Kawasaki-shi, Kanagawa-ken(JP)
Inventor: **Hirakawa, Hideki, Ohnabirasuzuki 202**
Chigasakiminami, Kohoku-ku
Yokohama-shi, Kanagawa-ken(JP)
Inventor: **Asahioka, Yoshimi**
3-6-23 Hiyoshi, Kohoku-ku
Yokohama-shi, Kanagawa-ken(JP)
Inventor: **Kimura, Mariko**
4-11-4 Himonya, Miguro-ku
Tokyo(JP)

(74) Representative: **Lehn, Werner, Dipl.-Ing. et al**
Hoffmann, Eitle & Partner Patentanwälte
Arabellastrasse 4
D-8000 München 81(DE)

(54) Method and apparatus for generating sentences.

(57) A method and an apparatus for generating sentences capable of generating more natural and easily comprehensible sentences by incorporating the selection of an appropriate choice for the form of expression or the order of words from a number of possible candidates in a process of sentence construction. In the apparatus, conceptual structures to be utilized as elements of sentences to be generated are stored; rules to be followed in generating sentences are stored; and sentences are generated from the stored conceptual structures in accordance with the stored rules by utilizing sizes of partial structures of the conceptual structures specified by the stored rules to determine a sentence structure of the sentence to be generated.

FIG.1



EP 0 387 876 A2

METHOD AND APPARATUS FOR GENERATING SENTENCES IN SUITABLE FORM OF EXPRESSION AND ORDER OF WORDS

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to a method and an apparatus for generating sentences from given conceptual structures as elements of the sentences.

Description of the Background Art

In a conventional apparatus for generating sentences, a form of expression and order of words in a sentence to be generated are uniquely determined once the conceptual structures for the sentence matches with a certain prescribed pattern. For this reason, in generating a sentence, it has conventionally been impossible to select a form of expression which is suitable from a point of view of a balance of the sentence as a whole, out of a number of different forms of expressions having equivalent meanings.

As an example, consider a case of generating an English sentence from a concept "..ing .... is good enough". In this case, two expressions "What is needed is to .. ...." and "..ing .... is sufficient" may be considered. Here, when a partial structure "...." is lengthy, the latter choice would result in a rather awkward sentence in which the subject is disproportionately large, so that the choice of the former would be preferable. On the contrary, when the partial structure "...." is short, the choice of the latter may be appropriate. However, this type of a choice of the form of expression has been impossible in the conventional apparatus for generating sentences.

As another example, consider a case of choosing an arrangement of the order of words. Namely, in English it is generally preferable to place a lengthy or complicated phrase toward the end of the sentence. Thus, for example a sentence "He had called the man on whose judgement he now had to rely an idiot." is rather awkward albeit grammatically correct. Alternate arrangement "He had called an idiot the man on whose judgement he now had to rely." would be less awkward and may b more preferable. Again, this typ of a choice of the ord r of words has been impossibl in the conventional apparatus for generating sentences.

It can b seen from these examples that, in ord r to be able to s lect an appropriate choice for the form of expression or the order of words from a number of possible candidates in a process of sentence construction, it is necessary to be able to assess a size of a partial structure of the sentence. However, the conventional apparatus for generating sentences has been impossible to incorporate such function, so that unnatural or awkward sentences have often been generated by the conventional apparatus for generating sentences.

SUMMARY OF THE INVENTION

It is therefore an object of the present invention to provide a method and an apparatus for generating sentences capable of generating more natural and easily comprehensible sentences by incorporating the selection of an appropriate choice for the form of expression or the order of words from a number of possible candidates in a process of sentence construction.

According to one aspect of the present invention there is provided an apparatus for generating sentences, comprising: conceptual structure storage means for storing conceptual structures to be utilized as elements of sentences to be generated; rule storage means for storing rules to be followed in generating sentences; and sentence generation means for generating sentences from the conceptual structures stored in the conceptual structure storage means in accordance with the rules stored in the rule storage means by utilizing sizes of partial structures of the conceptual structures specified by the rules stored in the rule storage means to determine a sentence structure of the sentence to be generated.

According to another aspect of the present invention there is provided a method of generating sentences, comprising the steps of: storing conceptual structures to be utilized as elements of sentences to be generated; storing rules to be followed in generating sentences; and generating sentences from the stored conceptual structures in accordance with the stored rules by utilizing sizes of partial structures of the conceptual structures specified by the stored rules to determine a sentence structure of the sentence to be generated.

Other features and advantages of the present invention will become apparent from the following description taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic block diagram of one embodiment of an apparatus for generating sentences according to the pr sent invention.

Fig. 2 is a schematic flow chart for the operation of the apparatus of Fig. 1.

Fig. 3 is a tree diagram for conceptual structures of an example of sentence to be generated by the apparatus of Fig. 1.

Fig. 4 is a tree diagram for conceptual structures of another example of sentence to be generated by the apparatus of Fig. 1.

Fig. 5 is an illustration of an example of rules for concepts to be used in the apparatus of Fig. 1.

Fig. 6 is a tree diagram for conceptual structures of still another example of sentence to be generated by the apparatus of Fig. 1.

Fig. 7 is an illustration of an example of rules for sentence constructions to be used in the apparatus of Fig. 1.

Fig. 8 is a detailed flow chart for the operation of the apparatus of Fig. 1.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring now to Fig. 1, there is shown one embodiment of an apparatus for generating sentences according to the present invention.

In this embodiment, the apparatus comprises a conceptual structure storage unit 1 for storing conceptual structures to be utilized as elements of sentences to be generated, a rule storage unit 2 for storing rules to be followed in generating sentences which includes rules for individual concepts 2a and rules for sentence structure construction 2b, and a sentence generation unit 3 for generating sentences from the conceptual structures stored in the conceptual structure storage unit 1 in accordance with the rules stored in the rule storage unit 2. Contents of the conceptual structures and the rules will become apparent from the following description of examples of sentence generation.

In general, the sentence generation by this apparatus is performed in two steps of a flow chart of Fig. 2 as follows.

First, at a conceptual structure modification step 31, with respect to a particular conceptual structure relevant to the sentence to be generated which is stored in the conceptual structure storage unit 1, a sentence structure and words appropriate for a desirable form of expression for the particular conceptual structure relevant to the sentence to be generated are selected out by using the rules for individual concepts 2a, and the conceptual structure relevant to the sentence to be generated are modified according to the selected sentence structure and words.

Then, at a sentenc structure construction step 32, an order of words are determined for the modifi d conceptual structure resulting from the conceptual structure modification step 31 in accordance with the rules for sentence structure construction 2b, so as to obtain the sentence as an ordered series of words.

In both of these steps 31 and 32, sizes of particular partial structures are utilized as keys for the selection and determination to be performed, which will now be described in detail by using examples of sentence generation.

First, the conceptual structure modification step 31 will be explained by using the specific conceptual structures "Reading this book is good enough" and "Delivering this letter to his office by tomorrow is good enough", which are shown diagrammatically in Figs. 3 and 4, respectively. In Figs. 3 and 4, each encircled entity will be referred to as a node, and nodes are labelled by integers increasing from top to bottom.

Now suppose that the specific rules shown in Fig. 5 are given for the specific concept of "..ing .... is good enough". Fig. 5 shows two rules for this concept, where each rule comprises a pattern of modification to be applied appearing on a first line, a condition of applicability to be satisfied in applying this rule appearing on a second line, and actions to be taken when this rule is applied appearing on a third line and below. Thus, the rule 1 states that when the partial structure below a phrase "..ing .... is good enough" (node 1) has no more than four nodes, a specific sentence structure "..ing .... is sufficient" is selected and the conceptual structure will be modified by putting the partial structure below the node 1 in a verbal form, whereas the rule 2 states that when the partial structure below a phrase "..ing .... is good enough" (node 1) has more than four nodes, a specific sentence structure "what is needed is to.. ...." is selected and the conceptual structure will be modified by putting the partial structure below the node 1 in a infinitive form.

According to these rules, the conceptual structure "Reading this book is good enough" will be modified to "Reading this book is sufficient" since in this conceptual structure the partial structure below the node 1 has only three nodes "this", "book", and "read", so that the rule 1 applies. Similarly, the conceptual structure "Delivering this letter to his office by tomorrow is good enough" will be modified to "What is needed is to deliver this letter to his office by tomorrow" since in this conceptual structure the partial structure below the node 1 has s ven nodes "tomorrow", "by", "this", "letter", "his", "office", and "deliver", so that the rule 2 applies.

Next, the determination of the order of words at the sentence structure construction step 32 will be

explained by using the specific conceptual structure "We call the on-line automatic translation typing dialogue via communication satellite ATTD" diagrammatically shown in Fig. 6 which is assumed to have resulted from the conceptual structure modification step 31. Here "automatic translation typing dialogue" is regarded as a single node.

Now suppose the specific rule shown in Fig. 7 is given as the rule for sentence structure construction. This rule concerns with the order of the objective and objective complement and specifies that such an objective which is either of less than 8 for weight value (pw), or, without a relative clause of more than 6 for weight value and a prepositional phrase of more than 6 for weight value, will be placed before the objective complement, while an objective not of the types enumerated above will be placed after the objective complement. Here, the weight value represents not a number of nodes involved but another quantity characterizing the partial structures such as a number of words involved, a number of letters involved, a number of syllables involved, a number of stressed syllables involved, a number of phonemes involved, or other similar quantities.

As an example, when the weight value is obtained as the number of words, the objective of this sentence is 8 because the single node "automatic translation typing dialogue" contains four words so that this node contributes 4 for the weight value, while each of four other nodes below this node contribute 1 for the weight value. As a consequence, in this case, the objective is placed after the objective complement according to the rule of Fig. 7, so that the sentence "We call ATTD the on-line automatic translation typing dialogue via communication satellite" will be generated, rather than a less desirable sentence "We call the on-line automatic translation typing dialogue via communication satellite ATTD".

Thus, the operation of the sentence generation by the apparatus of Fig. 1 can be summarized in a flow chart of Fig. 8 as follows.

When the sentence is to be generated, a conceptual structure relevant to the sentence to be generated is given to the conceptual structure storage unit 1 at the step 311.

Then, the sentence generation unit 3 takes the rules for the individual concept involved in this conceptual structure from the rule storage unit 2 at the step 312, find one of the rules for which this conceptual structure satisfies the condition of applicability by determining the size of th partial structure specified in the rule and checking the condition of applicability specified in the rul at the step 313, and applies this rule to this conceptual structure so as to desirably modify this conceptual structure by prosecuting the actions specified by this rule at the step 314.

Then, aft r the modified conceptual structure is obtained, the sentence generation unit 3 takes the rules for sentence structure construction from th rule storage unit 2 at the step 321, and applies these rules to this modified conceptual structure so as to desirably order the words of this modified conceptual structure at the step 322.

As described, according to this embodiment, it is possible to provide a method and an apparatus for generating sentences capable of generating more natural and easily comprehensible sentences by incorporating the selection of an appropriate choice for the form of expression or the order of words from a number of possible candidates in a process of sentence construction, which is made possible in this embodiment by utilizing the size of the particular partial structures.

It is to be noted that although, in the above embodiment, the number of nodes and number of words are used as the sizes of the partial structures in the conceptual structure modification step 31 and the sentence structure construction step 32, respectively, the size of the partial structure in either step may be obtained as any quantity characterizing the partial structure such as a number of nodes involved, a number of words involved, a number of letters involved, a number of syllables involved, a number of stressed syllables involved, a number of phonemes involved, or other similar quantities.

It is also to be noted that the method and apparatus of the present invention may be adapted beneficially to a machine translation system, or a language processing system such as a summary generation system, a translation telephone system, or a qustion answering system.

Besides these, many modifications and variations of the above embodiments may be made without departing from the novel and advantageous features of the present invention. Accordingly, all such modifications and variations are intended to be included within the scope of the appended claims.

Reference signs in the claims are intended for better understanding and shall not limit the scope.

Claims

1. An apparatus for generating sentences, comprising:
conceptual structure storage means(1) for storing conceptual structures to be utilized as el ments of sentences to b generated;
rule storage means(2) for storing rules to be followed in generating sentences; and
sentence generation means(3) for generating sen-

tences from the conceptual structures stored in the conceptual structure storage means(1) in accordance with the rules stored in the rule storage means(2),
CHARACTERIZED IN THAT the sentence generation means (3) generates the sentences by utilizing sizes of partial structures of the conceptual structures specified by the rules stored in the rule storage means(2) to determine a sentence structure of each sentence to be generated.

2. The apparatus of claim 1, wherein the sizes of partial structures are given in terms of numbers of nodes contained in the partial structures.

3. The apparatus of claim 1, wherein the sizes of partial structures are given in terms of numbers of words contained in the partial structures.

4. The apparatus of claim 1, wherein the sizes of partial structures are given in terms of numbers of letters contained in the partial structures.

5. The apparatus of claim 1, wherein the sizes of partial structures are given in terms of numbers of syllables contained in the partial structures.

6. The apparatus of claim 1, wherein the sizes of partial structures are given in terms of numbers of pitched words contained in the partial structures.

7. The apparatus of claim 1, wherein the sentence structure to be determined is an appropriate form of expression for the sentence to be generated.

8. The apparatus of claim 7, wherein the rules stored in the rule storage means(2) includes rules for individual concept, each rule being specifying condition of applicability and actions to be prosecuted on the conceptual structures, and wherein the sentence generation means(3) generates each sentence, on a basis of the conceptual structure relevant to the sentence to be generated stored in the conceptual structure storage means(1), by taking the rules for individual concept involved in this conceptual structure from the rule storage means-(2), finding one of the rules for which this conceptual structure satisfies the condition of applicability by determining the size of the partial structure specified in the rule and checking the condition of applicability specified in the rule, and applying this rule to this conceptual structure so as to desirably modify this conceptual structure by prosecuting the actions specified by this rule.

9. The apparatus of claim 1, wherein the sentence structure to be determined is an appropriate order of words for the sentence to be generated.

10. The apparatus of claim 9, wherein the rules stored in the rule storage means(2) includes rules for sentence structure construction specifying an order in which words of th conceptual structures are to be arranged, and wherein the sent nce generation means(3) generates each sentence, on a basis of the conceptual structure relevant to the sentence to be generated stored in the conceptual structure storage m ans(1), by taking the rules for sent nce structure construction from the rule storage means(2), and applying these rules to this conceptual structure so as to desirably order the words of this modified conceptual structure.

11. A method of generating sentences, comprising the steps of:
storing conceptual structures to be utilized as elements of sentences to be generated;
storing rules to be followed in generating sentences; and
generating sentences from the stored conceptual structures in accordance with the stored rules,
CHARACTERIZED IN THAT the sentences are generated by utilizing sizes of partial structures of the conceptual structures specified by the stored rules to determine a sentence structure of each sentence to be generated(31, 32).

12. The method of claim 11, wherein the sizes of partial structures are given in terms of numbers of nodes contained in the partial structures.

13. The method of claim 11, wherein the sizes of partial structures are given in terms of numbers of words contained in the partial structures.

14. The method of claim 11, wherein the sizes of partial structures are given in terms of numbers of letters contained in the partial structures.

15. The method of claim 11, wherein the sizes of partial structures are given in terms of numbers of syllables contained in the partial structures.

16. The method of claim 11, wherein the sizes of partial structures are given in terms of numbers of pitched words contained in the partial structures.

17. The method of claim 11, wherein the sentence structure to be determined is an appropriate form of expression for the sentence to be generated.

18. The method of claim 17, wherein the stored rules includes rules for individual concept, each rule being specifying condition of applicability and actions to be prosecuted on the conceptual structures, and wherein each sentence is generated, on a basis of the stored conceptual structure relevant to the sentence to be generated(311), by taking the rules for individual concept involved in this conceptual structure(312), finding one of the rules for which this conceptual structure satisfies the condition of applicability by determining the size of the partial structure specified in the rule and checking the condition of applicability specified in the rule-(313), and applying this rule to this conceptual structure so as to desirably modify this conceptual structure by prosecuting the actions specified by this rule(314).

19. The method of claim 11, wherein the sentence structure to be determined is an appropriate order of words for th sentence to b generated.

20. The method of claim 19, wherein the stored rules includes rules for sentence structure construction specifying an order in which words of the conceptual structures are to be arranged, and wherein each sentence is generated, on a basis of the stored conceptual structure relevant to the sentence to be generated(311), by taking the rules for sentence structure construction(321), and applying these rules to this conceptual structure so as to desirably order the words of this conceptual structure(322).

FIG.1

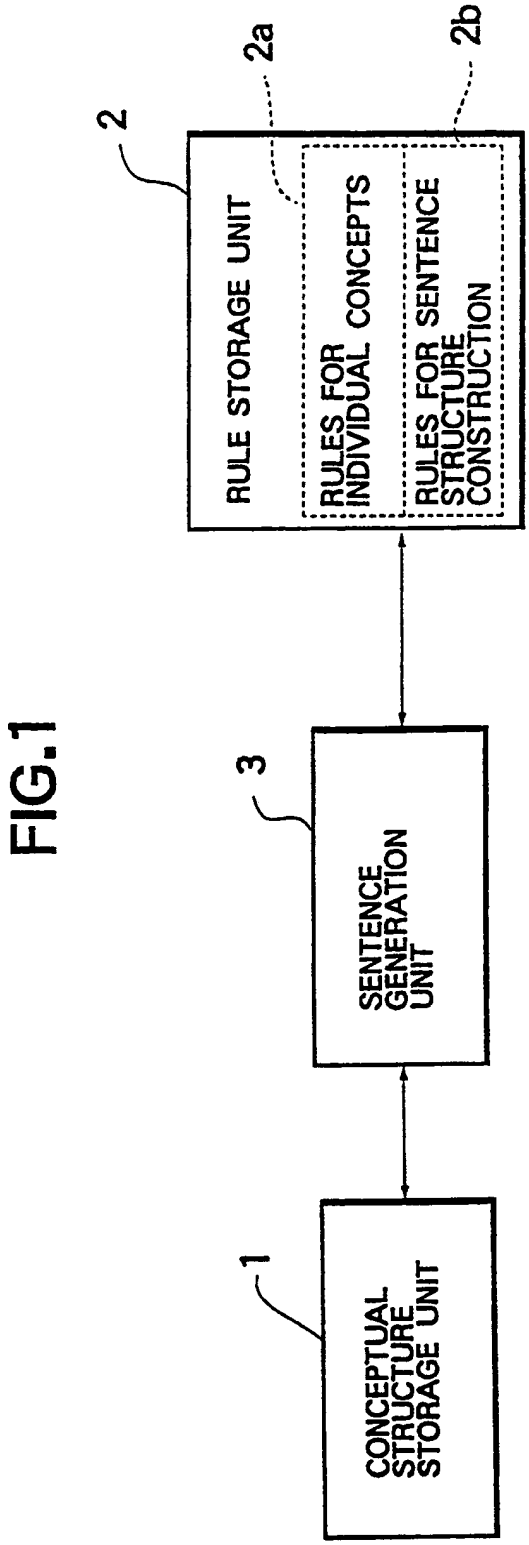

2
2a
2b
RULE STORAGE UNIT
RULES FOR INDIVIDUAL CONCEPTS
RULES FOR SENTENCE STRUCTURE CONSTRUCTION
3
SENTENCE GENERATION UNIT
1
CONCEPTUAL STRUCTURE STORAGE UNIT

FIG.2

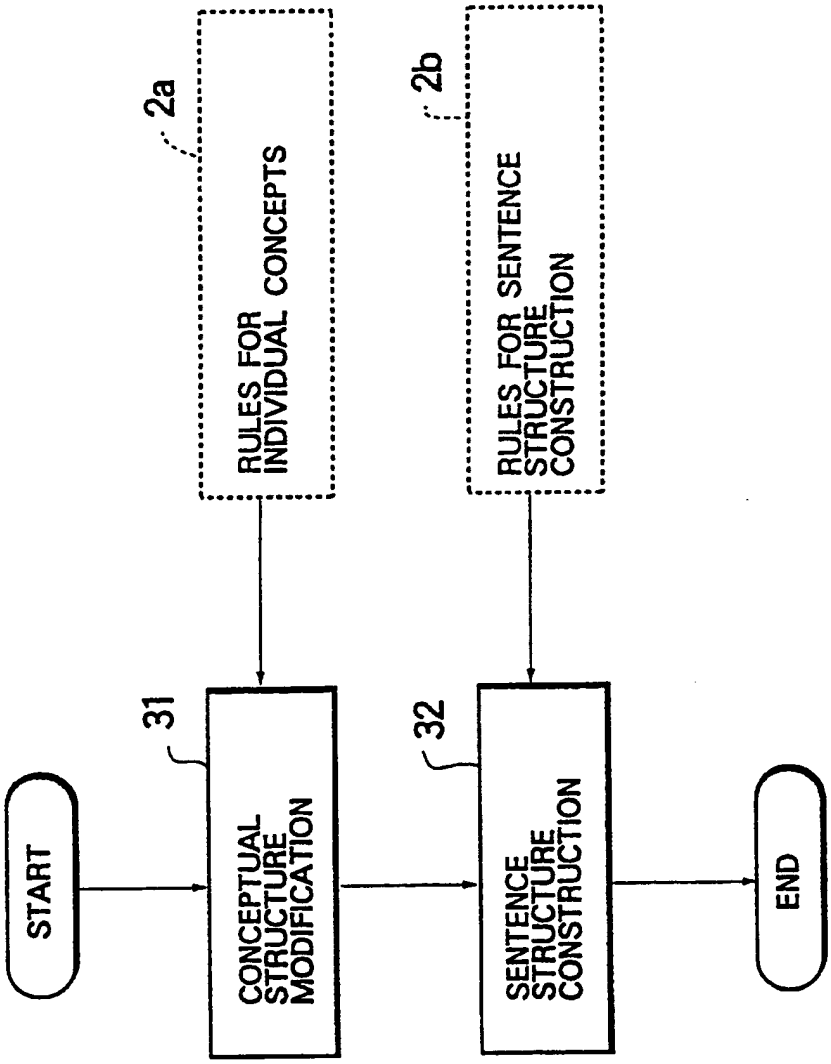

START
CONCEPTUAL STRUCTURE MODIFICATION
SENTENCE STRUCTURE CONSTRUCTION
END
RULES FOR INDIVIDUAL CONCEPTS
RULES FOR SENTENCE STRUCTURE CONSTRUCTION
31
32
2a
2b

FIG.3

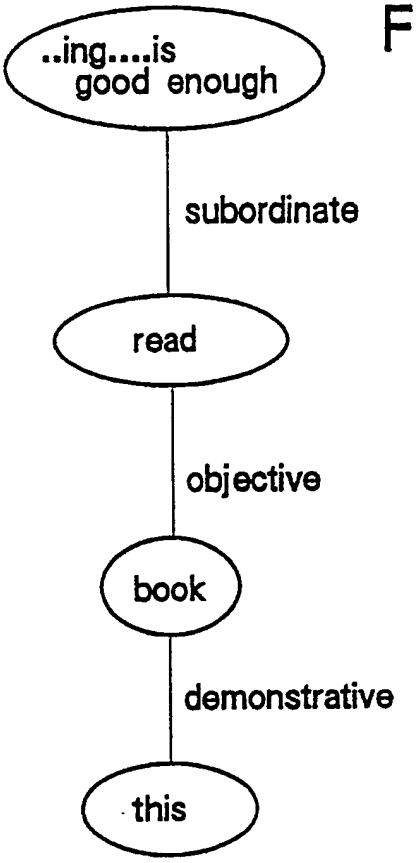

..ing....is
good enough
subordinate
read
objective
book
demonstrative
this

FIG.4

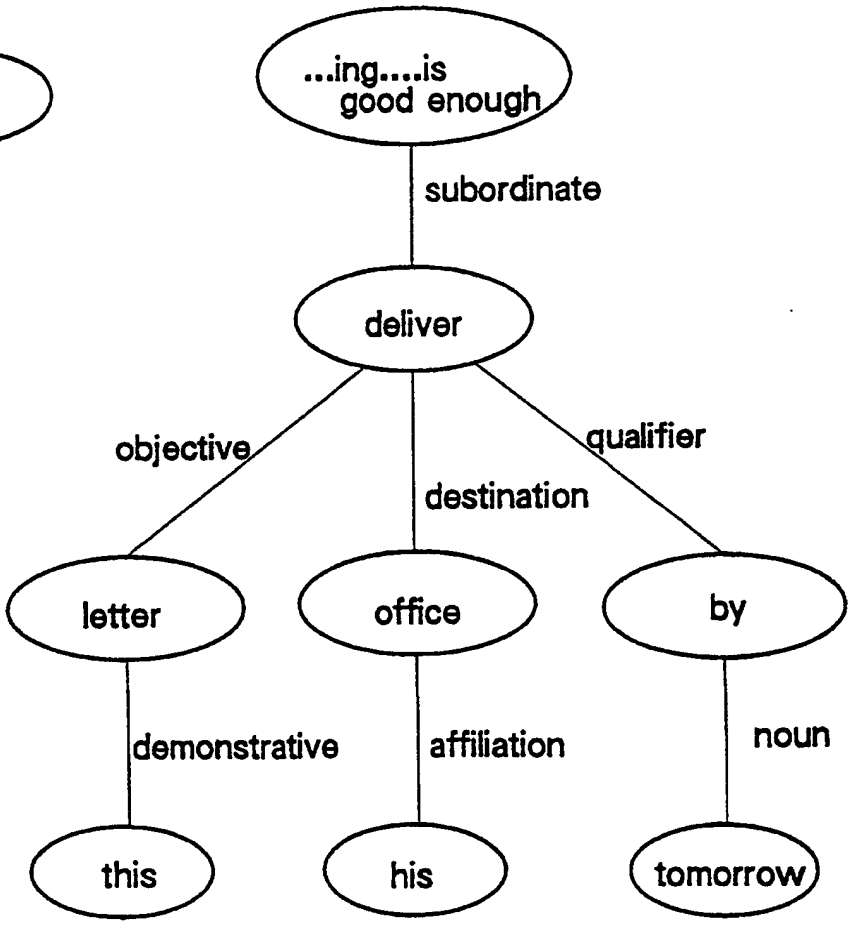

...ing....is
good enough
subordinate
deliver
objective
destination
qualifier
letter
office
by
demonstrative
affiliation
noun
this
his
tomorrow

FIG.5

	[...ing....is good enough]
1	Pattern: * (subordinate-1) ⇒ * (subject-1, subjective complement-2) Condition: the partial structure below the node 1 has no more than four nodes Actions: the partial structure below the node 1 is put in a verbal form a term "sufficient" is set for a new node 2 a term "be" is set for itself
2	Pattern: * (subordinate-1) ⇒ * (subject-1, subjective complement-2) Condition: the partial structure below the node 1 has more than four nodes Actions: the partial structure below the node 1 is put in an infinitive form a term "what is needed" is set for a new riode 2 a term "be" is set for itself

EP 0 387 876 A2

FIG.6

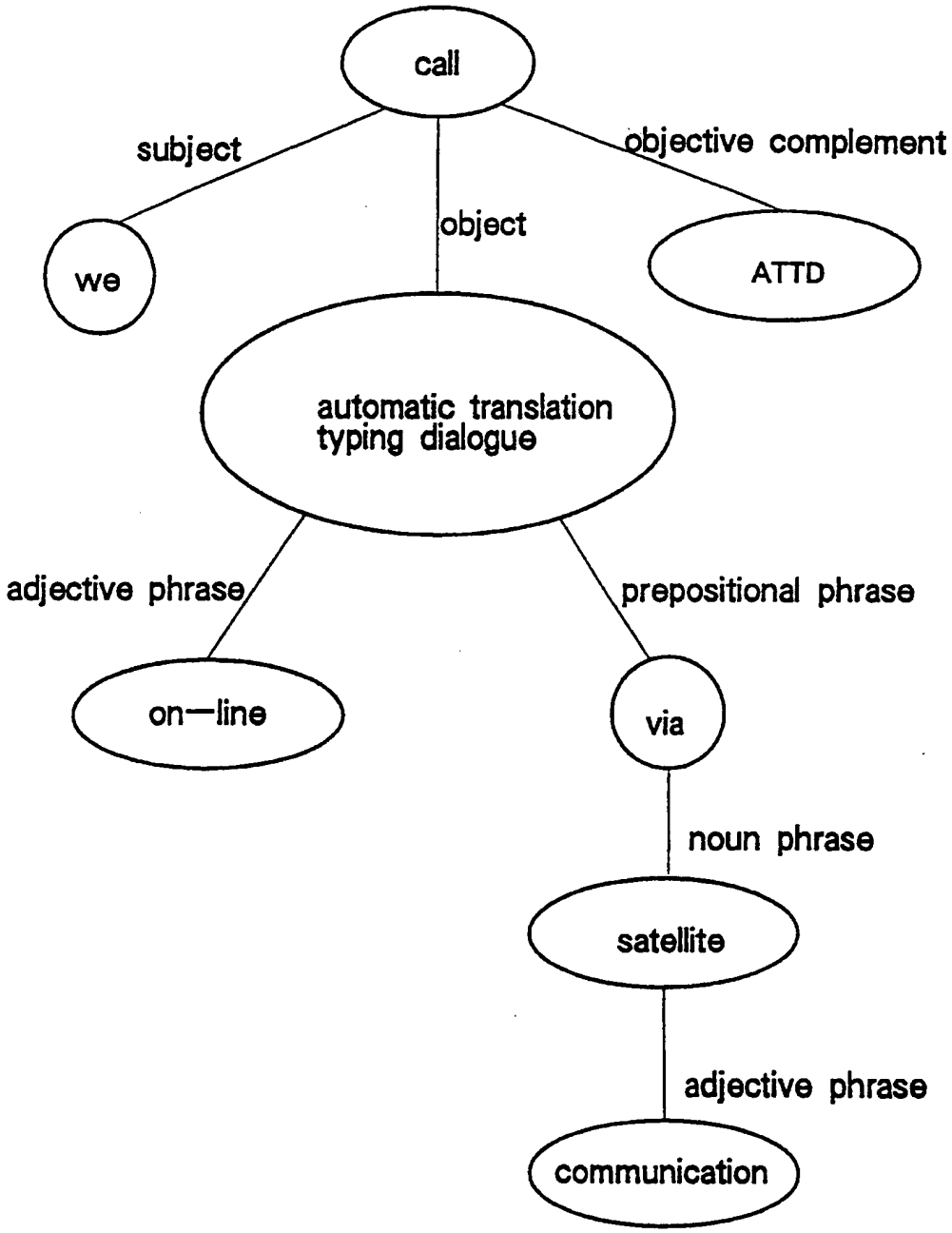
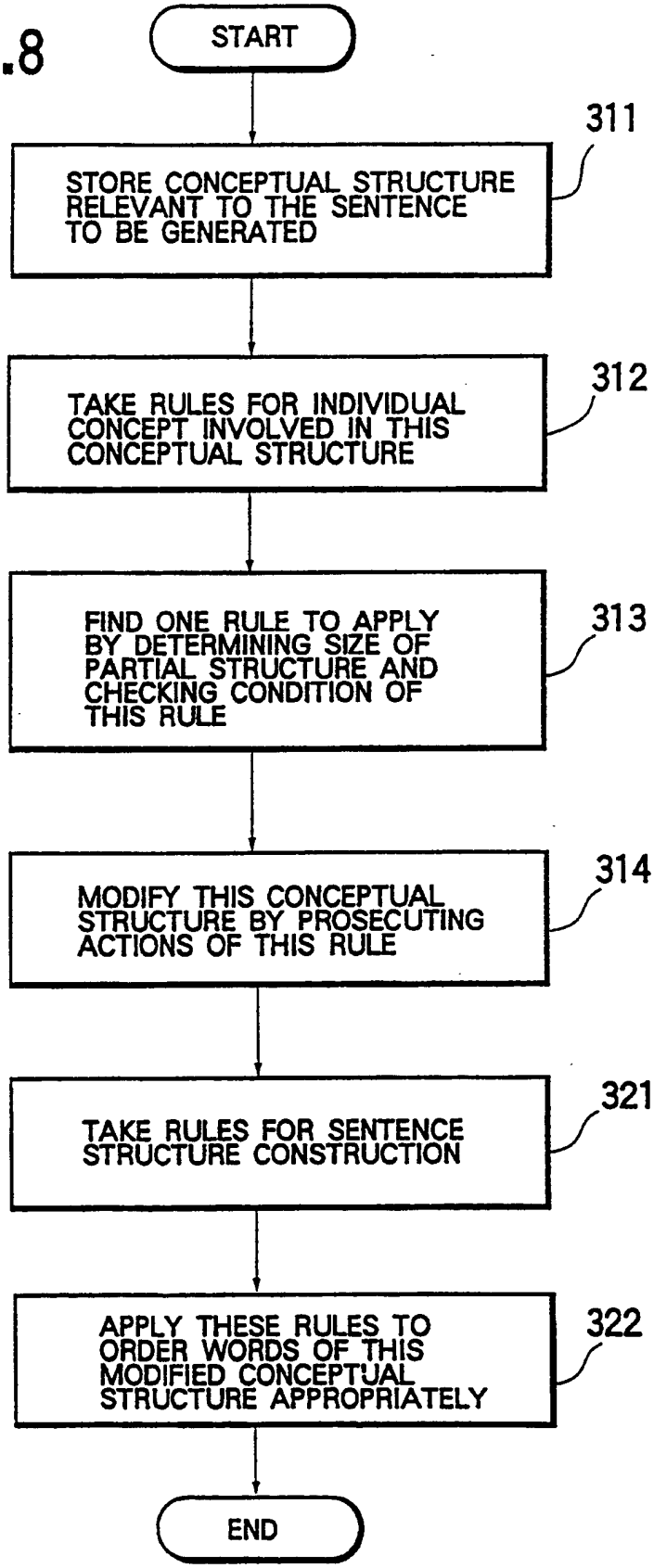

call
subject
object
objective complement
we
ATTD
automatic translation
typing dialogue
adjective phrase
prepositional phrase
on—line
via
noun phrase
satellite
adjective phrase
communication

FIG.7

```
. . .

objective {pw<8 | (without a relative clause of pw>6 & without a prepositional phrase of pw>6)} ,

objective complement,

objective,

. . .
```

FIG.8


START
311
STORE CONCEPTUAL STRUCTURE RELEVANT TO THE SENTENCE TO BE GENERATED
312
TAKE RULES FOR INDIVIDUAL CONCEPT INVOLVED IN THIS CONCEPTUAL STRUCTURE
313
FIND ONE RULE TO APPLY BY DETERMINING SIZE OF PARTIAL STRUCTURE AND CHECKING CONDITION OF THIS RULE
314
MODIFY THIS CONCEPTUAL STRUCTURE BY PROSECUTING ACTIONS OF THIS RULE
321
TAKE RULES FOR SENTENCE STRUCTURE CONSTRUCTION
322
APPLY THESE RULES TO ORDER WORDS OF THIS MODIFIED CONCEPTUAL STRUCTURE APPROPRIATELY
END